IJPCE
International Journal of Physics and Chemistry Education, 16(2), 21-36

www.ijpce.org
ISSN: 2589-8876

RESEARCH ARTICLE

A methodology for generating subquestions for the Force Concept Inventory (and other research-based assessments)

Michael M. Hull*
University of Alaska Fairbanks, Department of Physics

Jun-ichiro Yasuda
Yamagata University, Institute of Arts and Sciences

Naohiro Mae
Osaka University, Research Center for Nuclear Physics

Abstract

In this paper, we discuss the methodology we have developed and have been using for creating subquestions for the Force Concept Inventory (FCI). The FCI is a research-based assessment that is used internationally to assess student understanding of Newtonian mechanics. The assessment has been investigated from a number of perspectives and many suggestions have been made for its improvement. One challenge that is becoming more and more pervasive as more learning has transitioned to an online environment and more students are completing the FCI without a proctor is that of test security. If the answers to FCI items become easily accessible, then students will be able to provide correct answers despite lacking in understanding of Newtonian mechanics. One approach to mitigate the effects of items leaking into the public sphere and maintaining test security is the creation and administration of subquestions in place of the original FCI questions. Subquestions have an additional benefit of reducing false positives (answering a survey item correctly without correct understanding) and false negatives (answering incorrectly despite correct understanding). In this paper, we will discuss how we created subquestions for four items on the FCI, informed by survey-based interviews with students and the original intended targets of the items.

Keywords

Force Concept Inventory, subquestions, false positives, Newtonian mechanics

◆ Received 11 March 2023 ◆ Revised 1 June 2024 ◆ Accepted 20 June 2024

State of the literature

- The Force Concept Inventory (FCI) is a research-based assessment that is used on a wide scale to assess student understanding of Newtonian mechanics.
- Several items on the FCI are known to be prone to false positives (students answering the item correctly without correctly understanding the content), and false positives on all items are possible.
- The methodology of creating subquestions based upon false positives detected in survey validation interviews has been used to detect and quantify false positives.

Contribution of this paper to the literature

- We discuss subquestions we created for FCI items to enhance test security and allow for the reduction of false positives.
- We discuss in detail our methodology for creating these subquestions so that other researchers can create additional FCI subquestions.
- We discuss how this methodology can be used not only with the FCI, but with other research-based assessments as well.

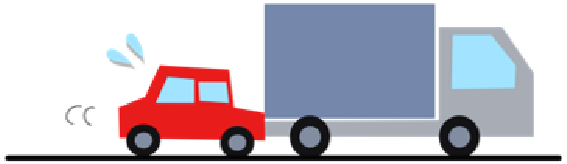
Introduction

The Force Concept Inventory (FCI) is a research-based assessment that has been and continues to be used internationally on a wide scale to assess student understanding of Newtonian mechanics (Hake, 1998). It is a 30-item multiple choice assessment, with each item requiring about one minute to complete; as such, the FCI requires most of a 50-minute class period to administer. The available responses on the survey items include distractors that are based upon common student conceptions that are documented in physics education research literature (Hestenes et al., 1992). Due in no small part to its wide-spread use, the FCI has been scrutinized from a number of perspectives and many suggestions have been made for its improvement. Using Rasch analysis, Planinic et al. (2010) concluded that, for the sample of high school students they investigated, the FCI is sufficiently unidimensional, but that there are several insufficiencies; namely, the difficulty of the test items is inappropriate for their student sample, there is a lack of easy items, and problems of middle difficulty are too similar to each other in terms of difficulty. Traxler et al. (2018) showed that some items in the FCI perform poorly in terms of gender fairness, indicating a need for replacement of these items with alternatives that are more fair. Scott et al. (2012) conducted a factor analysis and found that the intended item categories (as printed in the FCI taxonomy table) do not align well with student response patterns. For example, there was no factor found that is specific to kinematics.

In prior work, we considered false positives, which is when an item is answered correctly despite the respondent lacking understanding of the content (top-right quadrant in **Table 1**) (Yasuda et al., 2018; Yasuda & Taniguchi, 2013). Students' reasoning about individual FCI items is typically collected through survey validation interviews, where students are encouraged to think out loud

https://doi.org/10.51724/ijpce.v16i2.351
* Corresponding author

A large truck breaks down out on the road and receives a push back into town by a small compact car as shown in the figure below.



16. After the car reaches the constant cruising speed at which its driver wishes to push the truck:
A) the amount of force with which the car pushes on the truck is equal to that with which the truck pushes back on the car.
B) the amount of force with which the car pushes on the truck is smaller than that with which the truck pushes back on the car.
C) the amount of force with which the car pushes on the truck is greater than that with which the truck pushes back on the car.
D) the car's engine is running so the car pushes against the truck, but the truck's engine is not running so the truck cannot push back against the car. The truck is pushed forward simply because it is in the way of the car.
E) neither the car nor the truck exert any force on the other. The truck is pushed forward simply because it is in the way of the car.

Figure 1. Q.16 from the FCI, with the image redrawn. The correct answer is A.

while responding to each survey prompt while being recorded by the interviewer. False positives and false negatives (when an item is answered incorrectly despite understanding of the topic: bottom-left quadrant in **Table 1**) can also be determined by asking students to write down an explanation for each multiple choice response and then coding that free response for correctness. The free response prompt of each survey item can also be turned into a secondary multiple choice prompt, making the instrument two-tiered. For more information about how false positives and false negatives are identified, readers are referred to (Scott et al., 2012; Yasuda et al., 2018; Yasuda & Taniguchi, 2013; Hestenes & Halloun, 1995; Thornton et al., 2009; Wang & Bao, 2010; Yasuda et al., 2011, 2023; Taniguchi & Yasuda, 2014; Wilson & Low, 2015; Scott & Schumayer, 2017; Low & Wilson, 2017; Table I in Hestenes et al., 1992; Table II in Hestenes et al., 1992). Since items on the FCI are multiple choice with five possible responses to choose from, a random guess will result in a false positive 20% of the time. Although the distractors aim to reduce such random guessing and bring false positives below this 20% (Hestenes & Halloun, 1995), Hestenes et al. (1992) nevertheless found false positives to be "fairly common". Out of the 30 items on the assessment, Question 16 (Q.16) – which involves a car pushing a truck at constant speed – seems to be most prone to false positives (Scott et al., 2012; Yasuda & Taniguchi, 2013; Hestenes & Halloun, 1995; Thornton et al., 2009; Wang & Bao, 2010; Yasuda et al., 2011; Taniguchi & Yasuda, 2014; Wilson & Low, 2015; Scott & Schumayer, 2017; Low & Wilson, 2017) (see **Figure 1**). Students frequently respond to Q.16 with a false positive, correctly choosing response A but for a reason that is physically incorrect. Specifically, students argue that the forces are equal in size because the motion is uniform. In so doing, students show their conflation of the balance of forces on a single object with the equal and opposite forces on interacting objects: they are incorrectly applying Newton's first law, instead of thinking in terms of Newton's third law (Scott et al., 2012; Yasuda & Taniguchi, 2013; Thornton et al., 2009; Wang & Bao, 2010; Yasuda et al., 2011; Taniguchi & Yasuda, 2014; Wilson & Low, 2015; Scott & Schumayer, 2017; Low & Wilson, 2017).

Table 1. Contingency table of answers

		True attribute of a respondent	
		Understanding	**Not understanding**
FCI Question	Correct	True positive	False positive
	Incorrect	False negative	True negative

To detect these false positives and correct for the systematic error caused by those false positives (that is, a respondent's score being inflated due to false positives), Yasuda et al. (2018) developed subquestions for three FCI items that were frequently answered with a wrong reasoning- based false positive, Q.6, Q.7, and Q.16. They also created subquestions for Q.5, for which no reasoning-based false positives had been detected (since it is a multiple-choice question, false positives due to random guessing are still possible). Yasuda et al. (2018) judged that a correct answer is a true positive if the respondent correctly answered a corresponding set of subquestions (see also Yasuda & Taniguchi, 2013), and a false positive if any of those subquestions was answered incorrectly. Using this method, the score of a mid-level student was found to be inflated by about 10% from false positives on Q.6, Q.7, Q.16, and Q.5 alone (Yasuda et al., 2018). In consideration that false positives (and false negatives, for that matter) can manifest not only on these items but on any FCI item, we have been extending this work to develop subquestions for all 30 FCI items. These subquestions reduce the probability of false positive and false negative responses, since (in the case of false positives) the odds of answering not only the original FCI question but also the corresponding subquestions correctly while still not understanding the content are lower than the odds of just answering the original FCI question correctly.

In this paper, we will focus on Q.21, Q.22, Q.23, and Q.24 to demonstrate the methodology we developed and employed for creating subquestions for the FCI. Although we have discussed our process of creating subquestions in our prior work (Yasuda & Taniguchi, 2013), this was limited to subquestions created from false positives. This paper expands upon our prior work by discussing three additional methods for creating subquestions.

14. A bowling ball accidentally falls out of the cargo bay of an airliner as it flies along in a horizontal direction.
As observed by a person standing on the ground and viewing the plane as in the figure at right, which path would the bowling ball most closely follow after leaving the airplane?

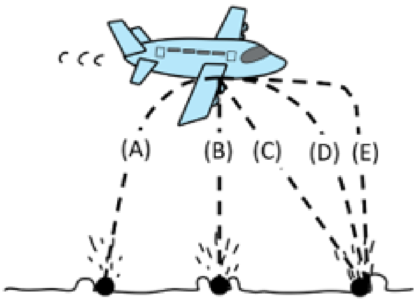


Figure 2. Q.14 from the FCI, with the image redrawn. The correct answer is D.

15. While the car, still pushing the truck, is speeding up to get up to cruising speed:
A) the amount of force with which the car pushes on the truck is equal to that with which the truck pushes back on the car.
B) the amount of force with which the car pushes on the truck is smaller than that with which the truck pushes back on the car.
C) the amount of force with which the car pushes on the truck is greater than that with which the truck pushes back on the car.
D) the car's engine is running so the car pushes against the truck, but the truck's engine is not running so the truck cannot push back against the car. The truck is pushed forward simply because it is in the way of the car.
E) neither the car nor the truck exert any force on the other. The truck is pushed forward simply because it is in the way of the car.

Figure 3. Q.15 from the FCI (the first part of the problem is the same as for Q.16 and is in Figure 1). The correct answer is A.

Methodology

Over the course of 3 months in 2010, one of the authors (JY) validated the Japanese translation of the FCI via conducting ten survey validation interviews (Yasuda et al., 2011) with Meijo University students (mostly engineering and science students) who had previously studied mechanics (Yasuda et al., 2011). As is typical in survey validation interviews, the interviewees completed the FCI item by item, thinking out loud as they did so. When it was unclear how the interviewee had arrived at a given answer, the interviewer asked follow-up questions. We noticed a number of false positives and false negatives in these interviews, and this served as our primary motivation for creating subquestions. Based upon these false positives and false negatives, we considered what subquestions might provide a more accurate measure of the true understanding of the respondents. Sometimes students answered incorrectly as a result of misreading a problem or not paying sufficient attention to a figure. For example, interviews revealed the potential for students to answer with a false negative on Q.14, involving a ball being dropped from a moving airplane (see **Figure 2**). Specifically, students can think about the problem from the perspective of someone riding in the airplane (in which case, the ball would seem to fall straight down) and to (incorrectly) choose path B (or, if air resistance is substantial, path A). A student who is reading sufficiently carefully should note the text "as observed by a person standing on the ground". If they look at the picture of the dirt for paths A and B, they should see that the ball could not have come from the left. Regardless, we made subquestions for this item because the point of the question is not to see how attentive to figure details or proficient at reading the student is. Rather, if it is feasible that students who understand projectile motion still get this item wrong, we want to make the problem even more straight-forward to ensure that their Newtonian understanding is being accurately measured.

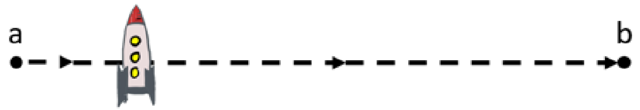
When making subquestions based upon the false positives and false negatives we detected, we refrained from introducing anything to the subquestions that was not motivated directly by the interviews. For example, we found that students can answer Q.15 (see **Figure 3**) with a false positive. Specifically, an interviewee chose the correct answer (that the forces are equal in magnitude) with the argument of *"I was thinking that, since the two vehicles are perfectly connected and continue at the same speed, the pushing force and the pushing back force are the same."* (All quotes from interviews were translated by the authors from Japanese into English). We consequently created a subquestion for Q.15 about a collision between bodies moving in opposing directions with different speeds to enable judgment of this false positive. We considered at first specifying that the faster object is also more massive, as our intuition was that doing so would allow us to catch yet additional false positives. However, since the interviewee had not discussed mass, we refrained from including information about mass in the subquestion.

For many of the FCI items, no false positives or false negatives were detected in the interview study. Since the interview study looked at only ten students, we considered it likely that a more intensive interview study might uncover additional false positives and false negatives. If nothing else, false positives can be caused on any item due to guessing, and we wished to more accurately quantify this effect. As such, we decided to create subquestions for all FCI items, although we would need tools beyond the interview study to accomplish this. In this paper, we detail our process of creating subquestions for Q.21, Q.22, Q.23, and Q.24, the four questions that constitute the block of items involving the motion of a rocket in space. A similar process was taken for the other FCI items. The four items are in **Figure 4–7**. A summary of the responses on Q.21-Q.24 from the ten interviewees is presented in **Table 2**.

Table 2. Codings of 10 interviewees (A – J) on Q.21, Q.22, Q.23, and Q.24. "TN" is an abbreviation for "true negative", "FP" is an abbreviation for "false positive", etc.

	A	B	C	D	E	F	G	H	I	J
Q.21	TN	FP	TN	TN	TN	FN	TP	TP	TP	TN
Q.22	TP	TN	TN	TN	TN	TP	TP	TP	TP	TN
Q.23	TN	FP	TN	TN	TP	TP	TP	TP	TP	FP
Q.24	TP	TP	TP	TP	TP	TP	TP	TP	TP	TP

A rocket drifts sideways in outer space from point "a" to point "b" as shown below. The rocket is subject to no outside forces. Starting at position "b", the rocket's engine is turned on and produces a constant thrust (force on the rocket) at right angles to the line "ab". The constant thrust is maintained until the rocket reaches a point "c" in space.



21. Which path below best represents the path of the rocket between points "b" and "c"?

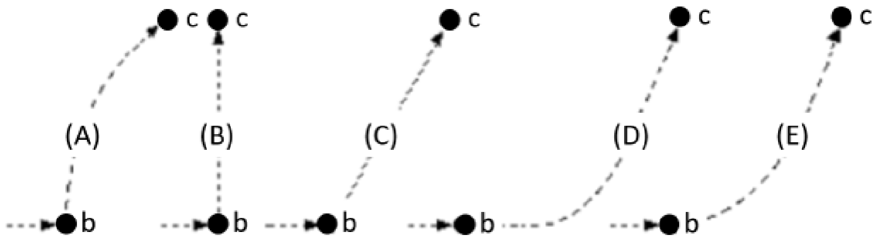


Figure 4. Q.21 from the FCI, with the figure redrawn. The correct answer is E.

22. As the rocket moves from position "b" to position "c", its speed is:
(A) constant.
(B) continuously increasing.
(C) continuously decreasing.
(D) increasing for a while and constant thereafter.
(E) constant for a while and decreasing thereafter.

Figure 5. Q.22 from the FCI. The correct answer is B.

23. At point "c" the rocket's engine is turned off and the thrust immediately drops to zero. Which of the paths below will the rocket follow beyond point "c"?

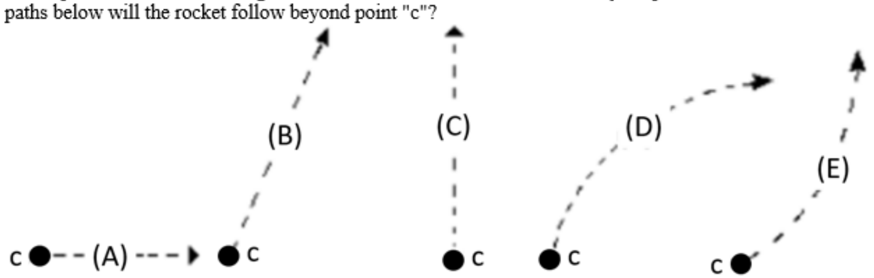


Figure 6. Q.23 from the FCI, with the figure redrawn. The correct answer is B.

24. Beyond position "c" the speed of the rocket is:
(A) constant.
(B) continuously increasing.
(C) continuously decreasing.
(D) increasing for a while and constant thereafter.
(E) constant for a while and decreasing thereafter.

Figure 7. Q.24 from the FCI. The correct answer is A.

On Q.22 and Q.24, there were no false negatives or false positives detected. Interviewee B exhibited a false positive (abbreviated "FP" in the table) on both Q.21 and Q.23. Interviewee F exhibited a false negative ("FN") on Q.21. Interviewee J exhibited a false positive on Q.23. Although these false positives and false negative allowed us a starting point for generating subquestions for Q.21 and Q.23, they were insufficient, as we will discuss below. To succeed in creating subquestions, we also utilized true positives from the interviews, the FCI taxonomy table (Table I in Hestenes et al., 1992), and the FCI misconceptions chart (Table II in Hestenes et al., 1992), as shown in **Figure 8**. We will discuss how each of these resources led to the creation of subquestions in turn.

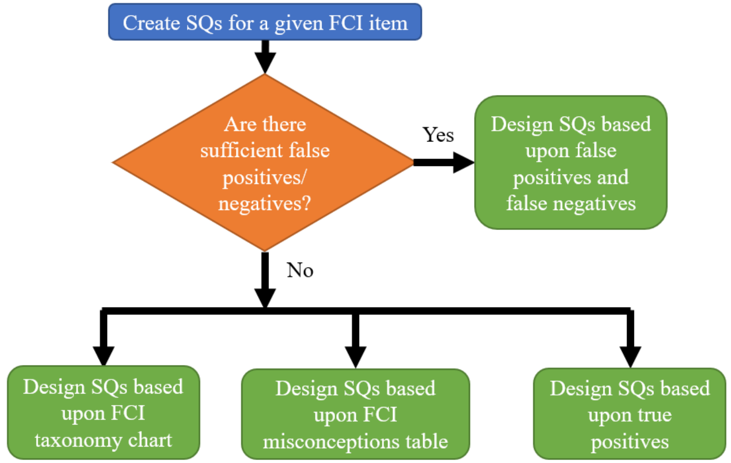


Figure 8. Flowchart showing overview of our methodology. Note that, although we demonstrate our methodology of creating subquestions (SQs) using the FCI as an example, it is relevant for research-based assessments in general: when false positives and false negatives are uncommon for a given survey item, consult the intentions of each survey item (here, "FCI taxonomy chart"), research on student difficulties leading to the instrument ("FCI misconceptions table") and true positives.

Resulting Subquestions from False Positives and False Negatives

Our first priority was to create subquestions that would be able to detect the false positives and false negatives we had discovered in the interviews. As mentioned above, interviewee B demonstrated false positives on both Q.21 and Q.23. For Q.21, interviewee B answered correctly but, when asked to explain, justified the response with "*since it's talking about outer space, I can't imagine what it would be like. Somehow, I just think this is the answer.*" An unsubstantiated response such as this is an example of a false positive due to guessing; unfortunately, such a response provides no hints for making subquestions.

Interviewee B's response to Q.23, on the other hand, was a different type of false positive. The student explained "*at first, it was going to the side, and when you add that to the upward force acting from the engine, I was thinking it would be like B.*" The justification involves adding a speed (which the interviewee perhaps associated to a "force of motion") and a force from the thrust. Although we recognize that it is possible that the interviewee meant that we should add the "going to the side" velocity component to the upwards velocity component *that results from* the engine's force acting for some time, this is putting words in the student's mouth. We decided to restrict our judgment to be based upon what the students said in the interviews (instead of our conjectures of what the students might have been thinking), and so we considered this response to be a false positive due to a lack of understanding. An explanation like what Interviewee B provided is what we would expect for a student who has the misconception "AF2: Motion implies active force" (Table II in Hestenes et al., 1992). This is consistent also with Interviewee B's response to Q.8. Another example of a false positive due to misunderstanding was interviewee J's response to Q.23: "*I was thinking that even if you turn off the engine, the force going forward is always acting in the same direction.*" This response describes a force acting even after the engine is turned off, suggesting a conflation between motion and force. We interpreted this statement to indicate that the student is thinking that if something is moving, there must be a force responsible. This reasoning is similar to the false positive from interviewee B, who also conflated motion and force in thinking that you can add the two together. These two false positives on Q.23 justified the creation of our first subquestion (SQ), shown in **Figure 9**.

The only false negative detected on these four items in the interview study was interviewee F's answer to Q.21, which resulted from finding the difference between D and E to be too subtle. We considered creating a subquestion for Q.21 that would include a caption to option D to specify that the first part of the trajectory is a horizontal line even though the engine is already turned on. In the end, however, we decided that the picture is sufficiently clear such that any such modification would more likely just make the picture harder to interpret.

Often related to false negatives, some interviewees offered advice on how to improve the FCI questions, either explicitly or implicitly by telling the interviewer what was confusing them about the problem statements. Interviewee B (who had false positives on Q.21 and Q.23) suggested

SQ23.1 A rocket drifts sideways in outer space from point "a" to point "b" as shown below. The rocket is subject to no outside forces (such as friction or air resistance). Starting at position "b", the rocket's engine is turned on and produces a constant thrust (force on the rocket) perpendicular to the line "ab". The constant thrust is maintained until the rocket reaches a point "c" in space.

At point "c" the rocket's engine is turned off and the thrust immediately drops to zero. After that, which of the following forces act on the rocket?

A) A force acting in the direction of motion
B) A force of inertia trying to make the rocket stop
C) A force acting in the direction the rocket was initially moving, in the direction of the line "ab"
D) A force acting in the direction the rocket was accelerated, perpendicular to the line "ab"
E) No forces act

Figure 9. SQ23.1, created based upon false positives detected in the interview study (the picture is the same as in Figure 4). The correct answer is E.

that it be explained what "outer space" and "thrust" mean. Similarly, interviewee E thought the set of space ship questions was the same as the set about the puck receiving a quick kick perpendicular to the direction of motion (Q.8 - Q.11), suggesting confusion about what "thrust" entails. However, since "thrust" is already explained in the problem statement as being a "force" and it is specified that this force is constant and applied for an extended period of time, we decided not to act on the suggestion to further explain "thrust". We did, however, act on the suggestion to specify that "no outside forces" includes air resistance.

Like with Q.23, it is often the case that study of false positives and false negatives is a good starting point for problem improvement and subquestion development, but is insufficient (only one subquestion was created in this case). In creating additional subquestions for Q.23, and for creating subquestions for Q.21, Q.22, Q.24, and other FCI items not discussed in this paper, we utilized additional resources. We will next discuss our usage of the FCI taxonomy table, which allowed us to see what the original FCI item was intended to assess.

Resulting Subquestions from the FCI Taxonomy Table and Misconceptions Chart

The FCI taxonomy table (Table I in Hestenes et al., 1992) classifies the correct responses of the FCI items into six groups of Newtonian concepts: Kinematics, First Law, Second Law, Third Law, Superposition Principle, and Kinds of Force. Additional sub-grouping within those groups specifies what each FCI item aims to assess. The taxonomy table tells us that Q.21 intends to probe student understanding about "Second Law: Constant force implies constant acceleration" and that "constant acceleration entails parabolic orbit". Similarly, Q.22 aims to see if students understand "Second Law: Constant force implies constant acceleration" and that "constant acceleration entails changing speed". On the other hand, Q.23 aims to probe student

SQ21.1 A rocket's engine produces a constant thrust (force on the rocket). External forces (friction, air resistance, etc.) acting on the rocket can be ignored. Which of the following best describes the acceleration of the rocket?
A) The acceleration is zero
B) The acceleration is constant, but non-zero
C) The acceleration constantly increases
D) The acceleration increases for some short amount of time and then remains constant
E) The acceleration increases for some short amount of time and then decreases

SQ21.2 A rocket drifts sideways in outer space from point "a" to point "b" as shown below. The rocket is subject to no outside forces (such as friction or air resistance). Starting at position "b", the rocket's engine is turned on and causes the rocket to have a constant acceleration at a right angle to the line "ab". The constant acceleration is maintained until the rocket reaches a point "c" in space.
Which path below best represents the path of the rocket between points "b" and "c"?

Figure 10. SQ21.1 and SQ21.2, created based upon the FCI taxonomy table. The correct answers are B and E, respectively. The figure for SQ21.2 is identical to that in Figure 4.

understanding about "First Law with no force… velocity direction constant". Finally, Q.24 similarly aims to probe student understanding about "First Law with no force… speed constant". These question goals give us insight into what modifications to make and what subquestions to develop. For example, we see that neither of these items intend to see if students understand that there is vacuum in outer space, and so we felt justified in expanding the description of "*the rocket is subject to no outside forces*" to specify that that includes air resistance, as discussed above. Furthermore, the subquestions we develop should aim to probe student understanding of the topics targeted by the corresponding original FCI items. This taxonomy table is particularly important for ensuring that the subquestions to each item measure what the original items intend, since factor analyses show that student response patterns deviate from the intended categories (e.g., Scott et al., 2012).

In the case of Q.21, we therefore created two subquestions which separated the ideas of 1) "constant force implies constant acceleration" and 2) "constant acceleration entails parabolic orbit". They are below in **Figure 10**.

Similarly, for Q.22, we aimed to separate the ideas of "Second Law: Constant force implies constant acceleration" and that "constant acceleration entails changing speed". Since the first of these two ideas is identical to that in Q.21, we found it appropriate to have SQ22.1 be identical to SQ21.1. The two resulting subquestions for Q.22 are in **Figure 11**.

As stated above, Q.23 aims to probe student understanding about "First Law with no force… velocity direction constant". We tentatively drafted a subquestion that would get at only this idea, which would use the same figure as in **Figure 6** and ask "*Two ice skaters are standing on a frozen lake and one pushes the other to the right. Which path best shows the motion of the ice skater after being pushed?*" In

SQ22.1) (Same as SQ21.1)

SQ22.2) A rocket drifts sideways in outer space from point "a" to point "b" as shown below. The rocket is subject to no outside forces (such as friction or air resistance). Starting at position "b", the rocket's engine is turned on and causes the rocket to have a constant acceleration at a right angle to the line "ab". The constant acceleration is maintained until the rocket reaches a point "c" in space.

As the rocket is traveling from point "b" to point "c", the speed of the rocket is

(A) constant.

(B) continuously increasing.

(C) continuously decreasing.

(D) increasing for a while and constant thereafter.

(E) constant for a while and decreasing thereafter.

Figure 11. SQ22.1 (identical to SQ21.1, in Figure 10) and SQ22.2, created based upon the FCI taxonomy table. The figure for SQ22.2 is identical to that in Figure 4. The correct answers are both B.

comparison with the original Q.23, we recognized that this subquestion is significantly easier. It is hard to imagine any students getting such a trivial item wrong. In cases such as this, we referenced also the FCI misconceptions chart (Table II in Hestenes et al., 1992), which specifies what misconceptions might lie behind selection of each distractor. We saw that distractors of Q.23 are designed to catch the following misconceptions: "Loss/recovery of original impetus" (choosing options A or D: specifically, the original impetus is recovered in these options), "impetus dissipation" (option D), "gradual/delayed impetus build-up" (option E), and "last force to act determines motion" (option C). The ice skater subquestion we had drafted did not allow for similar misconceptions-indicating distractors to be worked in, and we decided to replace this subquestion with a more challenging one that would incorporate the same spirit of distractors as Q.23 while still only focusing on the intended target of the original FCI question. To accomplish this, we separated the horizontal and vertical motion to have a clear path between the taxonomy table (about "First Law with no force… velocity direction constant") and the result of that velocity being up and to the right. We thus created one subquestion asking about motion to the right and one subquestion asking about motion upwards, both requiring understanding of the first law. The resulting subquestions (SQ23.2 and SQ23.3) are in **Figure 12**. As a specific example of how the misconceptions chart influenced the subquestion creation, we note that option D of SQ23.3 would likely be chosen by students who selected option A or option D on Q.23, reasoning with the "recovery of original impetus" idea.

SQ23.2) A rocket drifts sideways in outer space from point “a" to point “b" as shown below. The rocket is subject to no outside forces (such as friction or air resistance). Starting at position “b", the rocket's engine is turned on and produces a constant thrust (force on the rocket) perpendicular to the line "ab". The constant thrust is maintained until the rocket reaches a point "c" in space.

At point "c" the rocket's engine is turned off and the thrust immediately drops to zero. The speed of the rocket in the direction along the line “ab” is

A) Unchanged, because the force on the rocket was not in the direction along the line “ab”.
B) Unchanged, because the rocket was already moving in that direction before the engine was turned on.
C) Increased, because the engine sped the rocket up.
D) Decreased, because some of the speed in the direction of the line “ab” turned into speed in the direction perpendicular to the line “ab”.
E) Decreased, because the inertia of the rocket tried to make the rocket stop.

SQ23.3) A rocket drifts sideways in outer space from point “a" to point “b" as shown below. The rocket is subject to no outside forces (such as friction or air resistance). Starting at position “b", the rocket's engine is turned on and produces a constant thrust (force on the rocket) perpendicular to the line "ab". The constant thrust is maintained until the rocket reaches a point "c" in space.

At point "c" the rocket's engine is turned off and the thrust immediately drops to zero. After the engine is turned off, the speed of the rocket in the direction perpendicular to the line “ab”

A) remains constant.
B) continues to increase. Even after the rocket’s engine has been turned off, the rocket continues to feel the effect.
C) continues to increase at first, but after reaching a certain speed, it will become constant.
D) decreases because some of the speed in the direction perpendicular to the line “ab” will turn back into speed in the direction of the line “ab”.
E) decreases because of the rocket’s inertia that tries to make it stop.

Figure 12. Two subquestions for Q.23 based upon the FCI taxonomy table and misconceptions chart (the picture for each subquestion is the same as in Figure 4). The correct answer is A for both items.

The original FCI item Q.24 aims to probe student understanding about “First Law with no force… speed constant”. Both Q.23 and Q.24 introduce a complication beyond what the questions are aiming to assess (according to the taxonomy table) by asking about two-dimensional motion. With Q.23, trying to create a subquestion that is in only one dimension resulted in a subquestion that was too trivial to solve (see discussion above). Here in Q.24, however, we can consider a situation involving just one dimension without it becoming trivial, and two subquestions were created that are in **Figure 13**.

SQ24.1) Two ice skaters are standing on a frozen lake and one pushes the other. If external forces (friction, air resistance, etc.) can be ignored, then the speed of the skater after being pushed is…

(A) constant.

(B) continuously increasing.

(C) continuously decreasing.

(D) increasing for a while and constant thereafter.

(E) constant for a while and decreasing thereafter.

SQ24.2) Two ice skaters are standing on a frozen lake and one pushes the other. What is the horizontal force on the skater immediately after being pushed (when the skater is no longer in contact with the pusher)? Note that external forces (friction, air resistance, etc.) can be ignored.

A) The force of the skater’s inertia trying to make him stop

B) The skater carries the force of the pusher with him

C) The force of the skater’s motion in the direction of the push

D) The force of the skater’s motion in the direction opposite the push

E) There is no force on the skater

Figure 13. Two subquestions for Q.24 based upon the FCI taxonomy table. The correct answer for SQ24.1 is A and for SQ24.2 is E.

Our goal was to create 2-4 subquestions for each FCI question. When considering the false positives/negatives and the taxonomy chart were insufficient to accomplish this goal, we looked finally again at the interviews, but this time to see the reasoning employed by students coded with a true positive.

Resulting Subquestions from True Positives

Many times true positives are short one-step explanations like “a constant force means a constant acceleration.” Sometimes, however, more elaborate explanations are given that are comprised of a series of steps. Interviewee E correctly solved Q.22 with the following explanation: “If you break apart the velocity vector, the horizontal component is constant, but the vertical is slowly increasing, so overall it is slowly increasing.” We see in this explanation three cognitively irreducible components that came together to lead to the correct answer: 1) The horizontal component of velocity is constant, 2) The vertical component of velocity is increasing, 3) The

SQ22.3) A rocket drifts sideways in outer space from point “a" to point “b" as shown below. The rocket is subject to no outside forces (such as friction or air resistance). Starting at position “b", the rocket's engine is turned on and causes the rocket to have a constant acceleration at a right angle to the line “ab". The constant thrust is maintained until the rocket reaches a point "c" in space.

As the rocket is traveling from point “b” to point “c”, the speed of the rocket in the direction along the line “ab” is

A) Constant.
B) Continuously increasing.
C) Continuously decreasing.
D) Increases for some time and then becomes constant.
E) Is constant for some time and then decreases.

SQ22.4) A rocket drifts sideways in outer space from point “a" to point “b" as shown below. The rocket is subject to no outside forces (such as friction or air resistance). Starting at position “b", the rocket's engine is turned on and causes the rocket to have a constant acceleration at a right angle to the line “ab". The constant thrust is maintained until the rocket reaches a point "c" in space.

As the rocket is traveling from point “b” to point “c”, the speed of the rocket in the direction perpendicular to the line “ab” is

A) Constant.
B) Continuously increasing.
C) Continuously decreasing.
D) Increases for some time and then becomes constant.
E) Is constant for some time and then decreases.

Figure 14. Two subquestions for Q.22 based upon the true positive of interviewee E. The picture for each subquestion is the same as in Figure 4. The correct answer for SQ22.3 is A and for SQ22.4 is B.

sum of these two influences is a velocity that is increasing. Note that the intended assessment targets—the links between 1) constant force and constant acceleration, and 2) constant acceleration and changing speed—were essentially taken for granted in the student’s explanation. Rather, the student solved the problem with a reasoning chain not explicitly mentioned in the taxonomy table. We find it nevertheless plausible that, despite these three reasoning steps not being explicitly mentioned in the taxonomy table for this question, that the FCI developers would agree that these steps do reflect understanding of Newtonian mechanics and that they are indeed required to answer the question with a true positive. We hence attempted to turn each of these steps into additional subquestions. We struggled to find distractors for the final subquestion, which we envisioned as asking something like “If the horizontal component of velocity is a constant and the vertical component of velocity is increasing, then how does the velocity change?” and so settled upon the two subquestions in **Figure 14**. We reasoned that a student who correctly answers SQ22.3 and SQ22.4 would surely answer SQ22.5 correctly, unless the student had a false negative on the subquestion, and so we did not create SQ22.5.

Summary

In this paper, we have discussed our process for creating subquestions to the Force Concept Inventory (FCI). We looked in detail at Q.21 – Q.24, the item block concerning the motion of a rocket in space experiencing a temporary constant force in a direction perpendicular to the rocket’s initial motion. Our subquestion creation methodology began with detecting false positives and false negatives in interviews with students as they completed the FCI out loud. When an insufficient number of false positives and false negatives were found for a given item, we turned to other available tools. Specifically, we considered what the original FCI item was intended to assess, by consulting the FCI taxonomy chart (Table I in Hestenes et al., 1992). The multiple choice selections on the FCI include distractors aiming to judge if a student exhibits common student misconceptions. As such, we consulted also the misconceptions table (Table II in Hestenes et al., 1992) in the creation of our subquestions. Finally, to elucidate the individual steps required for correctly answering a given FCI item, we examined true positives of interviewees to see how they had broken down the problem, and then turned each step into a subquestion. A summary of the subquestions we created and the corresponding tool used in the creation of each subquestion is below in **Table 3**. Note that, although we demonstrated our methodology of creating subquestions using the FCI as an example, our methodology applies for creating subquestions on research-based assessments *in general*: when false positives and false negatives are uncommon for a given survey item, consult the intentions of each survey item, research on student difficulties leading to the instrument, and true positives.

Limitations and Future Work

As discussed earlier, our original and primary motivation for creating subquestions on the FCI was to reduce and quantify false positives and false negatives. Although many items on the FCI were answered by interviewees without false positives and false negatives, we can imagine that it is nevertheless possible for false positives and false negatives to arise from other respondents (at the very least, due to random guessing and being distracted, respectively). As we did in earlier research (Yasuda et al., 2018), we can calculate true positive ratios and true negative ratios for each FCI item by looking at student responses to the subquestions in comparison to responses on the corresponding original FCI items. We have not yet done this with the FCI items discussed in this paper, making this a limitation of our work thus far. Administering subquestions together with the original corresponding FCI item also allows for further validation of the subquestions.

Table 3. Summary of subquestions for Q.21-Q.24.

Original FCI question	Subquestion	Resource
Q.21) Which path below best represents the path of the rocket between points "b" and "c"?	SQ21.1) A rocket's engine produces a constant thrust … Which of the following best describes the acceleration … (Figure 10)	Taxonomy table
	SQ21.2) … constant acceleration… which path below best represents the path … (Figure 10)	Taxonomy table
Q.22) As the rocket moves from position "b" to position "c", its speed is:	SQ22.1) (Same as SQ21.1) (Figure 10)	Taxonomy table
	SQ22.2) …While the rocket is turned on and the acceleration is constant, the speed of the rocket is: (Figure 11)	Taxonomy table
	SQ22.3) As the rocket is traveling from point "b" to point "c", the speed… along the line "ab" is: (Figure 14)	True positive
	SQ22.4) As the rocket is traveling from point "b" to point "c", the speed… perpendicular to the line "ab" is: (Figure 14)	True positive
Q.23) At point "c" the rocket's engine is turned off and the thrust immediately drops to zero. Which of the paths below will the rocket follow beyond point "c"?	SQ23.1) … after the engine is turned off… what forces act on the rocket? (Figure 9)	False positive
	SQ23.2) … how does the horizontal component of the velocity before the engine was turned on compare with … after the engine was turned off? (Figure 12)	Taxonomy table; Misconceptions chart
	SQ23.3) After the engine is turned off… the vertical component of the velocity… (Figure 12)	Taxonomy table; Misconceptions chart
Q.24) Beyond position "c" the speed of the rocket is:	SQ24.1) Two ice skaters … one pushes the other… the speed of the skater after being pushed is… (Figure 13)	Taxonomy table
	SQ24.2) … What is the horizontal force on the skater immediately after being pushed … (Figure 13)	Taxonomy table

Although an additional limitation of our work is that we have not yet interviewed students to validate the prompts discussed in this paper, we have validated the subquestions of Q.8, which are pretty much identical to those of Q.23, through interviews with two students. For the most part, student responses were consistent in nature between the items, suggesting good functioning of the new subquestions. For example, on the original FCI question, the second interviewee correctly answered that the object would continue moving in a straight line at an angle to the original trajectory. The interviewee then answered two subquestions correctly, with arguments of "…adding in a perpendicular vector does not affect the other when it is divided into components" to the subquestion about the component of velocity in the original direction and "because there is no friction, no horizontal external force is applied" in responding to the subquestion about the component of velocity in the direction of the force (after the force is removed).

The creation and administration of subquestions has additional benefits as well. The FCI (and standardized exams in general) face the issue of items becoming over-exposed and the test becoming compromised, particularly as it becomes more common for the exam to be administered online. In an online asynchronous physics course taught by the first author, for example, a number of students who otherwise performed rather poorly in the class answered all 30 items of the FCI correctly (despite the typical introduction that extra credit is awarded just for completion, not for correctness, that the survey is meant to evaluate the instruction by comparing with their pretest scores, and that students should work alone as though the survey were a quiz!) An expanded item bank for the FCI is necessary, and subquestions can serve this purpose as substitutes for the original FCI items. We plan to integrate the subquestions we created into the item pool used by the computerized adaptive testing environment for the FCI (FCI-CAT) (Yasuda et al., 2021, 2022). In CAT, student ability is estimated with fewer questions administered (and hence in a shorter amount of time) by updating the student's ability estimate with each item that the student completes and avoiding items that are too easy or too difficult for the respondent.

We are also using subquestions in our research by administering the FCI items together with the corresponding subquestions in a Testlet format, where a given FCI item and its corresponding subquestions are administered together (Wainer et al., 2007). We can use Testlets in an online multi-stage testing environment, where respondents can review their item responses within each Testlet before moving on to the next Testlet. Some assessments, like the Lawson Test for Scientific Reasoning (Lawson, 1978) have multiple tiers (for example, one tier for the answer and another tier for the reasoning). One might administer the subquestions together with the original FCI item in a similar manner, using the subquestions as secondary tiers. A question to explore in future research is whether a partial credit model, where reduced points are awarded if a respondent answers a given subquestion incorrectly despite answering the original FCI item correctly, better estimates student ability. Although our previous work has judged such a case to be a false positive (see **Table 1**), the decision to award partial credit in this way is also justified, since it is possible that the incorrect response to the subquestion may have been a *false negative on the subquestion*.

A major limitation of our work is that we have only examined systematic error from false positives and false negatives from one population. Future work using subquestions to look at false positives and false negatives should also investigate whether the systematic error arising from false positives and false negatives is population-dependent. In the case of Q.16, for example, we expect that the false positive of answering in terms of forces cancelling out producing the constant speed would be much less common among extremely novice physics learners.

Acknowledgment

The authors thank Shizuka Nakayama for the artwork shown in Fig. 1, Fig. 2, and Fig. 4.

Disclosure of Interest

The authors have no competing interests to declare.

References

Hake, R. R. (1998). Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses. *American Journal of Physics*, *66*(1), 64-74. https://doi.org/10.1119/1.18809

Hestenes, D., & Halloun, I. (1995). Interpreting the force concept inventory: A response to March 1995 critique by Huffman and Heller. *The Physics Teacher*, *33*(8), 502. https://doi.org/10.1119/1.2344278

Hestenes, D., Wells, M., & Swackhamer, G. (1992). Force Concept Inventory. *The Physics Teacher*, *30*(3), 141-158. https://doi.org/10.1119/1.2343497

Lawson, A. E. (1978). The development and validation of a classroom test of formal reasoning. *Journal of Research in Science Teaching*, *15*(1), 11-24. https://doi.org/10.1002/tea.3660150103

Low, D. J., & Wilson, K. F. (2017). The role of competing knowledge structures in undermining learning: Newton's second and third laws. *American Journal of Physics*, *85*(1), 54-65. https://doi.org/10.1119/1.4972041

Planinic, M., Ivanjek, L., & Susac, A. (2010). Rasch model based analysis of the Force Concept Inventory. *Physical Review Special Topics - Physics Education Research*, *6*(1), Article 010103. https://doi.org/10.1103/PhysRevSTPER.6.010103

Scott, T. F., & Schumayer, D. (2017). Conceptual coherence of non-Newtonian worldviews in Force Concept Inventory data. *Physical Review Physics Education Research*, *13*(1), Article 010126. https://doi.org/10.1103/PhysRevPhysEducRes.13.010126

Scott, T. F., Schumayer, D., & Gray, A. R. (2012). Exploratory factor analysis of a Force Concept Inventory data set. *Physical Review Special Topics - Physics Education Research*. *8*(2), Article 020105. https://doi.org/10.1103/PhysRevSTPER.8.020105

Taniguchi, M.-a., & Yasuda, J.-i. (2014). Quantitative validation of Japanese translation of Force Concept Inventory using subquestions [Japanese]. *Journal of the Physics Education Society of Japan*, *62*(4), 226-231. https://doi.org/10.20653/pesj.62.4_226

Thornton, R. K., Kuhl, D., Cummings, K., & Marx, J. (2009). Comparing the force and motion conceptual evaluation and the force concept inventory. *Physical Review Special Topics - Physics Education Research*, *5*(1), Article 010105. https://doi.org/10.1103/PhysRevSTPER.5.010105

Traxler, A., Henderson, R., Stewart, J., Stewart, G., Papak, A., & Lindell, R. (2018). Gender fairness within the Force Concept Inventory. *Physical Review Physics Education Research*, *14*(1), Article 010103. https://doi.org/10.1103/PhysRevPhysEducRes.14.010103

Wainer, H., Bradlow, E. T., & Wang, X. (2007). What's a testlet and why do we need them? In *Testlet response theory and its applications* (pp. 44–59). Cambridge University Press. https://doi.org/10.1017/CBO9780511618765.005

Wang, J., & Bao, L. (2010). Analyzing force concept inventory with item response theory. *American Journal of Physics*, *78*(10), 1064-1070. https://doi.org/10.1119/1.3443565

Wilson, K. F., & Low, D. J. (2015). "On second thoughts…": Changes of mind as an indication of competing knowledge structures. *American Journal of Physics*, *83*(9), 802-808. https://doi.org/10.1119/1.4928131

Yasuda J.-i, Hull, M. M., & Mae, N. (2023). Visualizing depth of student conceptual understanding using subquestions and alluvial diagrams. *Physical Review Physics Education Research*, *19*(2), Article 020121. https://doi.org/10.1103/PhysRevPhysEducRes.19.020121

Yasuda, J.-i., & Taniguchi, M.-a. (2013). Validating two questions in the Force Concept Inventory with subquestions. *Physical Review Special Topics - Physics Education Research*, *9*(1), Article 010113. https://doi.org/10.1103/PhysRevSTPER.9.010113

Yasuda, J.-i., Hull, M. M., & Mae, N. (2022). Improving test security and efficiency of computerized adaptive testing for the Force Concept Inventory. *Physical Review Physics Education Research*, *18*(1), Article 010112. https://doi.org/10.1103/PhysRevPhysEducRes.18.010112

Yasuda, J.-i., Mae, N., Hull, M. M., & Taniguchi, M.-a. (2018). Analyzing false positives of four questions in the Force Concept Inventory. *Physical Review Physics Education Research*, *14*(1), Article 010112. https://doi.org/10.1103/PhysRevPhysEducRes.14.010112

Yasuda, J.-i., Mae, N., Hull, M. M., & Taniguchi, M.-a. (2021). Optimizing the length of computerized adaptive testing for the Force Concept Inventory. *Physical Review Physics Education Research*, *17*(1), Article 010115. https://doi.org/10.1103/PhysRevPhysEducRes.17.010115

Yasuda, J.-i., Uematsu, H., & Nitta, H. (2011). Validating a Japanese version of Force Concept Inventory [Japanese]. *Journal of the Physics Education Society of Japan*, *59*(2), 90-95. https://doi.org/10.20653/pesj.59.2_90

Appendix 1: List of all subquestions

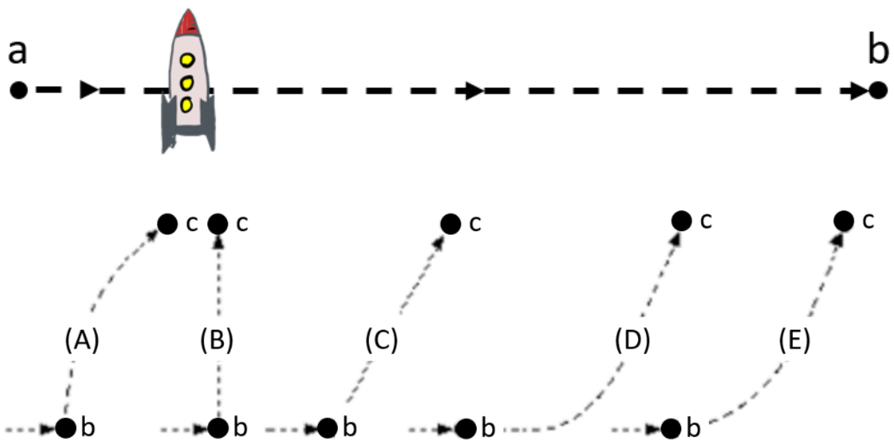
SQ21.1

A rocket’s engine produces a constant thrust (force on the rocket). External forces (friction, air resistance, etc.) acting on the rocket can be ignored. Which of the following best describes the acceleration of the rocket?

A) The acceleration is zero
B) The acceleration is constant, but non-zero
C) The acceleration constantly increases
D) The acceleration increases for some short amount of time and then remains constant
E) The acceleration increases for some short amount of time and then decreases

SQ21.2

A rocket drifts sideways in outer space from point “a" to point “b” as shown below. The rocket is subject to no outside forces (such as friction or air resistance). Starting at position “b”, the rocket’s engine is turned on and causes the rocket to have a constant acceleration at a right angle to the line “ab”. The constant acceleration is maintained until the rocket reaches a point “c” in space.



Which path below best represents the path of the rocket between points “b” and “c”?

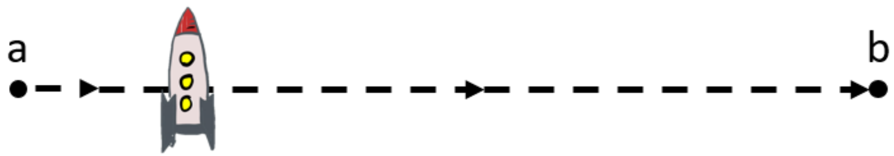
SQ22.1

A rocket’s engine produces a constant thrust (force on the rocket). External forces (friction, air resistance, etc.) acting on the rocket can be ignored. Which of the following best describes the acceleration of the rocket?

A) The acceleration is zero
B) The acceleration is constant, but non-zero
C) The acceleration constantly increases
D) The acceleration increases for some short amount of time and then remains constant
E) The acceleration increases for some short amount of time and then decreases

SQ22.2

A rocket drifts sideways in outer space from point “a” to point “b” as shown below. The rocket is subject to no outside forces (such as friction or air resistance). Starting at position “b”, the rocket’s engine is turned on and causes the rocket to have a constant acceleration at a right angle to the line “ab”. The constant acceleration is maintained until the rocket reaches a point “c” in space.

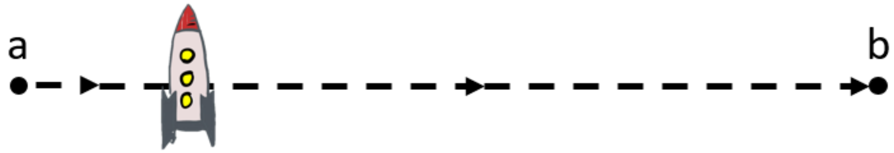


As the rocket is traveling from point “b” to point “c”, the speed of the rocket is

A) constant.
B) continuously increasing.
C) continuously decreasing.
D) increasing for a while and constant thereafter.
E) constant for a while and decreasing thereafter.

SQ22.3

A rocket drifts sideways in outer space from point “a” to point “b” as shown below. The rocket is subject to no outside forces (such as friction or air resistance). Starting at position “b”, the rocket’s engine is turned on and causes the rocket to have a constant acceleration at a right angle to the line “ab”. The constant thrust is maintained until the rocket reaches a point “c” in space.

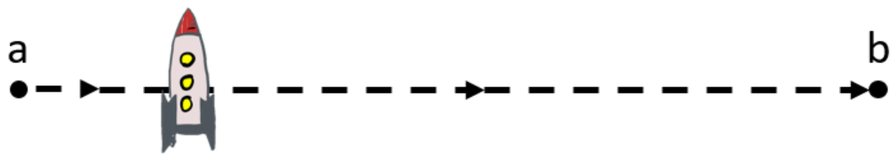


As the rocket is traveling from point “b” to point “c”, the speed of the rocket in the direction along the line “ab” is

A) Constant.
B) Continuously increasing.
C) Continuously decreasing.
D) Increases for some time and then becomes constant.
E) Is constant for some time and then decreases.

SQ22.4

A rocket drifts sideways in outer space from point “a" to point “b” as shown below. The rocket is subject to no outside forces (such as friction or air resistance). Starting at position “b”, the rocket”s engine is turned on and causes the rocket to have a constant acceleration at a right angle to the line “ab”. The constant thrust is maintained until the rocket reaches a point “c” in space.

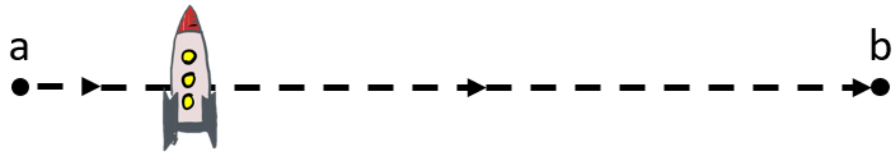


As the rocket is traveling from point “b” to point “c”, the speed of the rocket in the direction perpendicular to the line “ab” is

A) Constant.
B) Continuously increasing.
C) Continuously decreasing.
D) Increases for some time and then becomes constant.
E) Is constant for some time and then decreases.

SQ23.1

A rocket drifts sideways in outer space from point “a” to point “b” as shown below. The rocket is subject to no outside forces (such as friction or air resistance). Starting at position “b”, the rocket’s engine is turned on and produces a constant thrust (force on the rocket) perpendicular to the line “ab”. The constant thrust is maintained until the rocket reaches a point “c” in space.

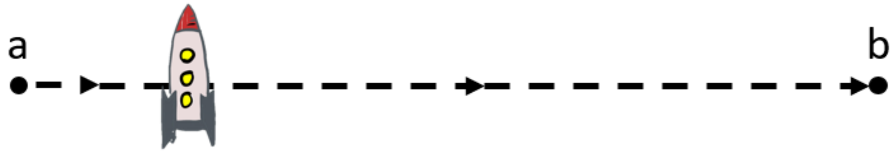


At point “c” the rocket’s engine is turned off and the thrust immediately drops to zero. After that, which of the following forces act on the rocket?

A) A force acting in the direction of motion
B) A force of inertia trying to make the rocket stop
C) A force acting in the direction the rocket was initially moving, in the direction of the line “ab”
D) A force acting in the direction the rocket was accelerated, perpendicular to the line “ab”
E) No forces act

SQ23.2

A rocket drifts sideways in outer space from point "a" to point "b" as shown below. The rocket is subject to no outside forces (such as friction or air resistance). Starting at position "b", the rocket's engine is turned on and produces a constant thrust (force on the rocket) perpendicular to the line "ab". The constant thrust is maintained until the rocket reaches a point "c" in space.

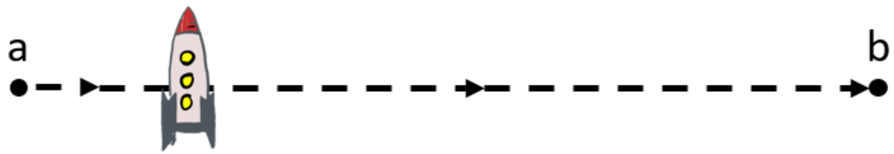


At point "c" the rocket's engine is turned off and the thrust immediately drops to zero. The speed of the rocket in the direction along the line "ab" is

A) Unchanged, because the force on the rocket was not in the direction along the line "ab".
B) Unchanged, because the rocket was already moving in that direction before the engine was turned on.
C) Increased, because the engine sped the rocket up.
D) Decreased, because some of the speed in the direction of the line "ab" turned into speed in the direction perpendicular to the line "ab".
E) Decreased, because the inertia of the rocket tried to make the rocket stop.

SQ23.3

A rocket drifts sideways in outer space from point "a" to point "b" as shown below. The rocket is subject to no outside forces (such as friction or air resistance). Starting at position "b", the rocket's engine is turned on and produces a constant thrust (force on the rocket) perpendicular to the line "ab". The constant thrust is maintained until the rocket reaches a point "c" in space.



At point "c" the rocket's engine is turned off and the thrust immediately drops to zero. After the engine is turned off, the speed of the rocket in the direction perpendicular to the line "ab"

A) remains constant.
B) continues to increase. Even after the rocket's engine has been turned off, the rocket continues to feel the effect.
C) continues to increase at first, but after reaching a certain speed, it will become constant.
D) decreases because some of the speed in the direction perpendicular to the line "ab" will turn back into speed in the direction of the line "ab".
E) decreases because of the rocket's inertia that tries to make it stop.

SQ24.1

Two ice skaters are standing on a frozen lake and one pushes the other. If external forces (friction, air resistance, etc.) can be ignored, then the speed of the skater after being pushed is…

A) constant.
B) continuously increasing.
C) continuously decreasing.
D) increasing for a while and constant thereafter.
E) constant for a while and decreasing thereafter.

SQ24.2

Two ice skaters are standing on a frozen lake and one pushes the other. What is the horizontal force on the skater immediately after being pushed (when the skater is no longer in contact with the pusher)? Note that external forces (friction, air resistance, etc.) can be ignored.

A) The force of the skater's inertia trying to make him stop
B) The skater carries the force of the pusher with him
C) The force of the skater's motion in the direction of the push
D) The force of the skater's motion in the direction opposite the push
E) There is no force on the skater

Appendix 2: Interviewee responses to Q.21 – Q.24

In this appendix, we present transcript and our coding decision for interviewee A on FCI items Q.21-Q.24 as an example of our analysis that led to **Table 2**. Transcripts from the other 9 interviews are available upon request.

The first interviewee ("Interviewee A"), answered Q.21 incorrectly, choosing option B, with the explanation of "*I was torn between options B and C, since there is no force from a to b. But, from a to b it is constant speed, and from b to c it is constant acceleration, so if you think about it carefully, it might be E. When I chose B, I was just thinking about the force vector.*"

At first, we considered this transcript to suggest that Interviewee A had a solid grasp on the physics necessary to answer Q.21 correctly, but nevertheless did not do so, and that this should thus be coded as a "false negative", resulting from the interviewee failing to carefully read that the arrows are not forces, but rather "paths". After further discussion, however, we decided that such an explanation is unlikely. Were it the case that the interviewee were just misreading what the arrows represent, the interviewee would have thought that that there is right-pointing force from a to b, since the path is drawn like that as well. Most likely, we concluded, this is a true negative, and that is the code we settled upon.

On Q.22, Interviewee A correctly chose option B, with the explanation of "the force is constant, so the acceleration is constant". This is a correct answer and it is accompanied by evidence of correct conceptual understanding, so we coded the response as a "true positive".

On Q.23, Interviewee A incorrectly chose option D, with the explanation of "because even if the engine is turned off, I think it will continue to move a little bit. Even in space it will decelerate, so I chose D. If it does not decelerate, it would be B." The interviewee was correct to say that if there is no deceleration, then B would be the correct answer, but this was not the final answer of the interviewee. D is the incorrect path, even if there would be a deceleration in space (it would still be path B). As such, we coded this as a "true negative".

Finally, on Q.24, Interviewee A correctly (but in contradiction with the answer to Q.23) chose option A with the explanation that "in space, things do not slow down". This was coded as a true positive.